AD_____

Award Number: DAMD17-01-1-0112

TITLE: Genetic Epidemiology of Prostate Cancer

PRINCIPAL INVESTIGATOR: Susan L. Neuhausen, Ph.D.

CONTRACTING ORGANIZATION: University of California, Irvine
Irvine CA 92697-7600

REPORT DATE: March 2007

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-03-2007	Final	01 Jan 01 – 28 Feb 07

4. TITLE AND SUBTITLE
Genetic Epidemiology of Prostate Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-01-1-0112

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Susan L. Neuhausen, Ph.D.

E-Mail: sneuhaus@uci.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of California, Irvine
Irvine CA 92697-7600

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT:
Prostate cancer results from complex interactions among genetic, endocrine, and environmental factors. Understanding genetic risk factors that contribute to the occurrence of prostate cancer is crucial to design both preventative and therapeutic strategies and to identify at-risk individuals. Plausible candidates for susceptibility genes for prostate cancer risk include genes involved in insulin-like growth factor signaling, androgen signaling, and in immune response. We hypothesized that genetic variation in genes in these pathways was associated with prostate cancer risk. We studied 199 incident prostate cancer cases and 263 age-matched controls. Genotyping was performed for 59 polymorphisms in 18 genes and statistical analyses performed to look at their associations with prostate cancer risk and aggressiveness. As previously reported, the IRS1 G972R GR/ RR genotypes (rs1801278) were associated with a significant 2.7-fold increased risk for prostate cancer (95% CI 1.5-4.9, p=0.0007). Other significant results were SNP rs2139924 in IGF1R with a 2.6-fold increased risk (95% CI 1.1-6.5) and rs361072 in PI3KCB with a 1.6 fold increased risk (95% CI 1.03-1.66). These results support a role of the insulin-like growth factor pathway in the etiology of prostate cancer. These results need to be replicatedin multiple, larger studies.

15. SUBJECT TERMS
prostate cancer, genetics, molecular variants, genetic epidemiology

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	8	USAMRMC

Table of Contents

Grant DAMD17-01-1-0112
Period: June 1, 2001-February 28, 2007

INTRODUCTION

Prostate cancer is the most commonly diagnosed cancer in men and the most common cause of cancer mortality in the United States (American Cancer Society, 2006). The American Cancer Society estimated that, during 2006, approximately 234,460 new cases of prostate cancer would be diagnosed in the United States. One in six men will be diagnosed with prostate cancer over the course of a lifetime. In our aging population, research leading to a reduction in the incidence of and mortality from prostate cancer is an urgent necessity. A critical problem in prostate cancer is an understanding of risk factors involved in disease development and aggressiveness. Clinically important genetic risk factors that may result in differences in individual susceptibility to prostate cancer likely include genes involved in androgen biosynthesis, metabolism and regulation and in prostate cell growth and death (Papatsoris et al., 2005; Schaid, 2004). We propose to answer the following questions. What are some of the genetic risk factors that determine who develops prostate cancer? Of those individuals who develop cancer, what risk factors contribute to the age at diagnosis and to aggressiveness of the disease? Using a case-control design, we are testing the hypothesis that common genetic polymorphisms (variants) in genes directly and indirectly involved in altering hormonal levels and prostate cell growth are associated with prostate cancer risk. We are investigating their associations with occurrence of prostate cancer, age at diagnosis, and aggressiveness of the disease as measured by Gleason score and tumor stage-related variables.

BODY

Our progress is described by Tasks. We were granted a second no-cost extension to complete the statistical analyses and to genotype and analyze additional SNPs in genes that were either in the IGF pathway or that had been described in the literature. We studied 199 prostate cancer cases and 263 age-matched controls. Data on family history, age at diagnosis, and clinical and pathological characteristics were obtained for the prostate cancer cases. We conducted genotyping for the genetic variants in DNA samples from this set of cases and controls.

Aim 1: To assay samples for the genetic variants (genotyping).

Task 1: Design allele-specific primers for genotyping. Test and optimize the genetic assays. Compare with published protocol results. Sequence to confirm that detecting the appropriate alleles. COMPLETED.

We performed the single nucleotide polymorphism (SNP) genotyping using the MGB Taqman exonuclease (Applied Biosystems) and the MGB Eclipse probe (Nanogen) assays. We were able to obtain assays for all 51 SNPs to be genotyped in Aim 5, including assays for 3 SNPs in *IGFBP1*, 4 SNPs in *IGFBP3*, 8 SNPs in *IGF1*, 18 SNPs in *IGF1R*, 3 SNPs in *IRS1*, 4 SNPs in *SHBG*, 4 SNPs in IL-1RN, and single SNPs for *PI3KCB* -359 T>C, *SHC1* M300V, *SRD5A2* V89L, *CDH1* rs16260, *TLR4* rs11536889, and *IL-10* rs1800896. For design of the assays, we first submitted sequence encompassing the SNP to the Applied Biosystems (ABI) Assay-by-Design service for designing the assays. For those that ABI could not design, we sent sequence to Nanogen (formerly Epoch Biosciences). They use the MGB Eclipse probe assay, which has a very different probe design than the Taqman assay. The majority of SNPs that could not be

designed by Assay-by-Design were able to be designed by Epoch. We then tested and optimized the assay when we received it using known homozygotes and heterozygotes for the variants. For the *IGFBP3* -202 A>C SNP, neither assay could be designed so we performed a restriction endonuclease assay.

Task 2: Screen for variants in insulin-like growth factor binding protein I (*IGFBP-1*) to identify a variant(s) for genotyping. COMPLETED

We identified 19 variants of which 3 SNPs are required to tag the variation in this gene in Caucasians. These three haplotype-tagging SNPs were used for genotyping in Task 5.

Task 3: Identify male controls which match prostate cancer cases. COMPLETED.

For the 199 prostate cancer cases for whom we had DNA and diagnosis and follow-up data, we identified 263 age-matched male population-based controls. Dr. Brothman was no longer able to enroll participants in his study, so there were no additional cases or controls available for this study.

Task 4: Aliquot DNA from all samples available. COMPLETED.

In the last year of the study, we were running low on 60 of the DNA samples. Therefore, we performed whole genome amplification on those samples using a kit from Amersham in order to increase the amount of DNA. We had previously validated that the DNA after whole genome amplification provided the same results as the non-amplified DNA.

Task 5: Perform genotyping. COMPLETED

The total DNA samples available for genotyping were 462 including 199 prostate cancer cases and 263 controls. Early on, we completed genotyping on the 199 prostate cancer cases and 263 controls for the following polymorphisms of which *SRD5A2-str*, *VDR-polyA, IGF1 192,* and *CYP11A* STR are short-tandem repeat markers with multiple SNPS and were analyzed on the ABI3100 sequencer; and of which *CYP17-MspI, VDR-BsmI, VDR-Taq1, SHBG* D327N, the *INS* +1127 Ins-PstI,*IRS1* G972R, *IGFBP3* -202 A>C and *IRS2* G1079D were analyzed using restriction digests. The remaining 50 SNPs were genotyped using Taqman assays as described in Task 1. For all polymorphisms, we would re-genotype if there was more than 4% missing data.

Note: In a previous report, we described that we had decided that it was not useful to genotype the microsatellite repeat markers in *HSD3B2* and *HSD17B2*, as there was no indication that they would be related to function.

Task 6: Read genotypes and enter into database. COMPLETED

Genotypes were uploaded into excel spreadsheets after analyses on the ABI3100 for the STR SNPs and on the ABI7900HT for the SNPs and manually entered for SNPs genotyped by restriction digests.

Aim 2: To statistically analyze the association of genes assayed in Aim 1 with prostate cancer age at diagnosis and aggressiveness, as measured by Gleason score and tumor stage-

related variables. Aim 3: To statistically analyze the association of genes assayed from Aim 1 with occurrence of prostate cancer.

Task 7: Design data entry forms for entering data into Sybase. COMPLETED at University of Utah, but then did not use.

This task was completed so that we could download the data into Sybase. However, since it was a finite amount of data, we decided it was better to use Excel spreadsheets that are uploaded for analysis with the statistical package SAS. Therefore, we did not use Sybase for storing the data.

Task 8: Edit data. Add data from medical records and Utah Cancer Registry. COMPLETED.

The prostate cancer cases were diagnosed from 1992-2000. Characteristics of the cases and controls are shown in Table 1. Age at diagnosis ranged from 45-78 years with a mean age of 62.6 years and a median age of 63 years. Of the tumors, 10 were well-differentiated, 139 were moderately differentiated and 50 were poorly differentiated. Thirteen of the cases had another type of cancer, either previous to or after diagnosis of prostate cancer. We obtained follow-up data on these cases with the dates of last follow-up ranging from 2000-2002. These data are in the excel spreadsheet with the genotypes. Of the 199 prostate cancer cases, 15 are deceased including 1 case diagnosed at 49 years of age who died from metastatic prostate cancer.

Table 1. Description of the study population

	Cases (%)	Controls (%)
Total available	199	263
Mean age (age range) at diagnosis/enrollment	63 (45-78)	64 (35-79)
Gleason category		NA*
1	0 (0)	
2	103 (51.7)	
3	71 (35.7)	
4	25 (12.6)	
Stage		NA*
T1	26 (13.1)	
T2	100 (50.3)	
T3	56 (28.1)	
T4	15 (7.5)	

*Not applicable

Task 9: Months 25-27: Test models and analysis methodologies. COMPLETED.

Gleason scores were placed into groupings commonly used in clinical prognosis. Group 1 was Gleason 1-3 (none in this study); group 2 was Gleason 4-6; group 3 was Gleason 7; and group 4 was Gleason 8-10. Unconditional logistic regression models were used to assess the main effects of the genetic variants on occurrence of prostate cancer. Gene x gene interactions were analyzed by logistic regression using the Wald $\chi 2$ test to determine significant differences in slopes. Logistic regression for a polychotomous outcome was used to assess associations with Gleason score ($\leq$6, 7, and $\geq$8). Gene x gene interactions for Gleason score were not analyzed as there are too few individuals with Gleason scores 8-10 in order to reliably fit a model with interactions.

Since the majority of the population was non-Hispanic white, adjustment for racial group was not performed.

We used a haplotype-tagging approach to examine the genetic variation in *SHBG, IRS1, IGFBP1*, *IGFBP3, IGF1* and *IGF1R*. This allowed us to examine the genetic variation across the entire gene in order to not miss a possible association within the gene. There are additional steps to haplotype analysis that are not present in traditional, genotype-based case-control studies. First, a set of SNPs must be selected that will mark the common haplotypes in the population. These SNPs are commonly referred to as haplotype-tagging SNPs. The next step is the assignment of haplotypes to the case and control individuals, based on their haplotype-tagging SNP genotypes. Without genotype information in the parents or a direct molecular assay of individual chromosomes, the haplotypes must be assigned based on a probability model. We have developed algorithms for selecting haplotype-tagging SNPs and estimating haplotype assignments for the sampled individuals. The second algorithm assigns haplotypes to each individual, based on the individuals' genotype data and the estimated population haplotype frequencies. The output is a matrix with a column for each of the common haplotypes present in the study population (frequency greater than 0.05) and a row for each individual. A logistic regression can be carried out with the haplotype data to estimate the risk of disease associated with each haplotype.

Task 10: Months 26-36: Perform statistical analyses as outlined in Methods.

Statistical analysis as described in Task 9 was performed for all polymorphisms genotyped in Task 5 to investigate the association of the genetic variants and the outcomes of risk of prostate cancer, Gleason category, and tumor stage. Where there were multiple SNPs in genes, selected as haplotype-tagging SNPs, we also analyzed the haplotypes. Prior to investigating the outcomes, we first determined whether the polymorphisms were in Hardy-Weinberg equilibrium (HWE) and if not, we eliminated them from further analysis, a conservative approach. For individual polymorphisms, not haplotypes, we investigated, additive, dominant, and recessive models. Of all the polymorphisms analyzed, only single SNPs in *IGF1R, IRS1*, and *PI3KCB* were significantly associated with risk of prostate cancer (Table 2). There were no significant interactions between the *IGF1R, IRS1*, and *PI3KCB* SNPS.

Table 2. SNPs showing statistically significant association with risk of prostate cancer

				Additive	Dominant	Recessive
Gene		case	control	OR (95%CI)	OR (95%CI)	OR (95%CI)
IGF1R	rs2139924	181	261	1.14 (0.82-1.58)	0.98 (0.66-1.45)	**2.65 (1.09-6.46)**
IRS1	rs1801278	177	259	**2.43 (1.35-4.39)**	**2.41 (1.28-4.53)**	4.44 (0.46-43.11)
PI3KCB	rs361072	180	263	1.15 (0.88-1.51)	0.96 (0.64-1.43)	**1.65 (1.03-2.66)**

For Gleason category and AJCC tumor stage, only *IRS1* G972R (rs1801278) was associated with an increasing risk for higher Gleason category ($p = 0.001$) and for higher tumor stage ($p = 0.004$) (Neuhausen et al., 2005).

Task 11: Months 34-36: Prepare and submit final report and manuscripts.

A final manuscript is being prepared now describing the significant association with IGF1R.. In addition, we submitted an abstract for the DOD Prostate Cancer Meeting in fall, 2007.

KEY RESEARCH ACCOMPLISHMENTS:

- Investigated the association of 59 polymorphisms in 18 genes that have been implicated to be key plays in prostate cancer development
- Identified that the *IRS1* G972R GR/RR genotypes are associated with a 2.7-fold increased risk of prostate cancer risk, the
- Determined that the *IGF1R* rs2139924 SNP is associated with a 2.6-fold increased risk of prostate cancer
- Found that PI3KCB rs361072 SNP is associated with a 1.65-fold increased risk of prostate cancer

REPORTABLE OUTCOMES:
IRS1 results: Neuhausen S et al., 2005. Prostate cancer risk and IRS1, IRS2, IGF1, and INS polymorphism: strong association of IRS1 G972R variant and cancer risk. The Prostate 64:168-174.

IGF1R results: Chu, LH et al. 2007. Abstract submitted for 2007 DOD Impact meeting. A manuscript is in preparation.

CONCLUSIONS: The *IRS1* G972R GR/RR genotypes are associated with a 2.7-fold increased risk of prostate cancer risk, a SNP in intron 10 of *IGF1R* is associated with a 2.6-fold increased risk of prostate cancer, and a SNP in *PI3KCB* is associated with a 1.6-fold increased risk. These results provide additional support for an insulin-like growth factor signaling in the etiology of prostate cancer. IRS1 is one of the key docking proteins to start the signaling cascade through phosphorylation of the IGF1R. PI3KCB is one of the first proteins in the phosphoinositol-3-kinase (PI3K) pathway that is phosphorylated, a key pathway in apoptosis. More work is needed to identify what the causal SNPs are in both *IGF1R* and *PI3KCB*. Furthermore, these results need to be replicated in larger case-control studies.

REFERENCES:

American Cancer Society. 2006. Cancer Facts and Figures. Atlanta, GA: American Cancer Society.

Neuhausen SL, Slattery ML, Garner CP, Ding YC, Hoffman M, Brothman AR. 2005. Prostate cancer risk and IRS1, IRS2, IGF1, and INS polymorphism: strong association of IRS1 G972R variant and cancer risk. The Prostate 64:168-174.

Papatsoris A, Karamouzis MV, Papavassiliou AG. 2005. Novel insights into the implication of the IGF-1 network in prostate cancer. Trends Mol Med 11: 52-55.

Schaid D. 2004. The complex genetic epidemiology of prostate cancer. Hum Molec Genet 13, Review Issue 1: R103-R121.

APPENDICES: none